Review of *Fundamentals of Matrix Computations*
by David S. Watkins*

Stephen G. Nash
Operations Research and Applied Statistics Department
George Mason University
Fairfax, Virginia 22030

Submitted by Richard A. Brualdi

This textbook covers the most important topics in computational linear algebra, explains them in considerable detail, has many exercises, has pointers toward software, and describes methods suitable for various vector and parallel computers. It occupies a niche below that of *Matrix Computations* by G. H. Golub and C. Van Loan. Individuals who find the level of *Matrix Computations* challenging might find this book more accessible. The Preface indicates that it was written for "advanced undergraduates, graduate students, and mature scientists," and I would not disagree. Where I do disagree with the author is in the need to motivate the material and to make connections between the theoretical ideas and their practical consequences.

The topics covered are the solution of linear systems of equations, least-squares problems (both via the *QR* factorization and via the SVD), and eigenvalue problems. For linear equations, many variants of Gaussian elimination are described. These include the usual special cases of banded and positive-definite systems, as well as column-oriented and block algorithms suitable for novel computer architectures. There is a chapter on the sensitivity of linear systems. The least-squares and eigenvalue problems are treated with similar thoroughness, with extensive discussions of algorithmic alternatives and sensitivity analysis. The book does not come with software, nor do the exercises make reference to specific software packages. Instead, various exercises require the implementation of algorithms described in the text, and an appendix makes suggestions about software. The exercises explicitly mention the use of FORTRAN as a programming language, but any language could be used without much altering the material.

---

*John Wiley & Sons, New York, 1991, 449 pp.

655 Avenue of the Americas, New York, NY 10010 1992 0024-3795/92/$5.00

Even though this is a book about "computations," it is not a "methods" book. It is written in a mathematical style with theorems and proofs. The book is largely self-contained, with background theoretical material provided. Many of the exercises are questions about the theory.

This may be all that many readers will care about. The book contains a wide range of material, develops that material carefully, and has an awareness of computational issues. I was not fully satisfied with the book, but not for technical reasons. I was dissatisfied because I did not think that the book adequately explained why computational linear algebra is important. Let me discuss this in connection with the chapters on least-squares problems.

Data fitting is certainly an important application of computational linear algebra. It is used both informally to smooth and display data for graphical display and more formally to estimate parameters in models, for example by statisticians in the context of regression. The development of least squares as the most important data-fitting technique has an illuminating history.

Various approaches to data fitting were proposed in the eighteenth and early nineteenth century by such notable mathematicians as Euler, Lagrange, and Gauss. Euler's early attempts were failures because he was not able to find a satisfactory way to combine measurements made under different circumstances. Other scientists developed special techniques for estimating parameters, but they often depended on the data and model under study and did not generalize. More successful approaches were discovered when scientists focused on the *errors* in the model, and made assumptions about the form or distribution of the errors. This was not fully satisfactory, however, since there was no good reason to prefer one error distribution over another—as long as they both satisfied a few basic requirements such as symmetry about zero. Many proposals were made, one of the last being the use of least squares by Gauss and Lagrange—the formula e^{-x^2} not being the most obvious thing to consider as the basis for an error distribution. Least squares was attractive because it was computationally easy to apply, but there was no good justification for choosing least squares over the many other data-fitting techniques that had been suggested. Gauss had proposed an explanation, but he had used a circular argument that was unsatisfactory.

The final resolution came with the central limit theorem, proved by Laplace, but based on a development going back to Pascal more than a century before. The central limit theorem justified the assumption that the errors in the data were normally distributed. Not only was least squares a computationally easy technique, it was a mathematically justifiable one as well. It soon spread to many countries and many disciplines. Even the general public became aware of it, thanks to well-publicized studies of birth rates, death rates, suicide rates, conviction rates, and much else.

Watkins does not completely ignore this background when he discusses least squares. He describes a data-fitting problem (fitting a straight line to data). He then mentions that various norms of the residual vector could be minimized to produce a fit and then concludes: "By far the nicest theory is that based on the 2-norm, and it is that theory that we will study in this chapter." Watkins is not the only author to introduce least-squares data fitting in this brief way, so why do I think it important to mention its history in a book on matrix computations?

First, least-squares data fitting (regression) is a central topic in statistics, and there is a huge literature associated with it. (Watkins only includes one statistics book in the bibliography, and it is referenced in association with the computation of the angles between two subspaces.) Second, the central limit theorem and the normal distribution indicate when it is appropriate to use least-squares techniques and motivate other methods such as robust estimation when the assumptions underlying least squares are not satisfied. Third, computational difficulties associated with least squares can be closely tied to the data-fitting problem. For example, (near) rank deficiency can arise because of missing data, redundant terms in the model, duplication of data points, or deficiencies in sampling. Since regression is so widely used in statistical applications, and since many of the conclusions that statisticians would like to draw from regression models depend crucially on the assumption that the model is properly specified (i.e., full rank), it is of great practical importance to make explicit connections between results in numerical analysis and their application in the context of data fitting.

I have concentrated here on the chapters on least-squares data fitting, but the question of motivation arose throughout my reading of the book (although the material on eigenvalue problems discusses applications at greater length). If the student or the reader is already aware of the significance of the material, then there would be no reason to be unhappy with the book. Also, for use in courses at a higher level, it might be reasonable to assume that such motivating material would be unnecessary. The book is otherwise effective and worthy of serious consideration.

Received 22 July 1991; final manuscript accepted 7 August 1991